AIP Conference Proceedings

On the use of cluster analysis for interpretation of soil pollution monitoring data obtained near mining enterprise

Cite as: AIP Conference Proceedings **2048**, 060016 (2018); https://doi.org/10.1063/1.5082131
Published Online: 11 December 2018

Maxim A. Medvedev, Alexander N. Medvedev, and Anna Kolomytseva

View Online

Export Citation

ARTICLES YOU MAY BE INTERESTED IN

About some economic applications of cohort analysis
AIP Conference Proceedings **2048**, 060015 (2018); https://doi.org/10.1063/1.5082130

AIP | Conference Proceedings
Get 30% off all print proceedings!
Enter Promotion Code PDF30 at checkout

AIP Publishing

© 2018 Author(s).

On the Use of Cluster Analysis for Interpretation of Soil Pollution Monitoring Data Obtained Near Mining Enterprise

Maxim A. Medvedev[1,2, a)], Alexander N. Medvedev[1,2, b)], Anna Kolomytseva[3)]

[1] *Ural Federal University, Mira st., 19, Yekaterinburg, Russia*
[2] *Institute of Industrial Ecology UB RAS, Sofia Kovalevskaya st., 20, Yekaterinburg, Russia*
[3] *Donetsk National Technical University, Artem 58, Donetsk, Ukraine, 83015*

[a)] Corresponding author: medvedevmaa@gmail.com
[b)] medvedev@ecko.uran.ru

Abstract: Assessment of soil contamination in areas affected by mining facilities is a necessary part of the research during the local environmental monitoring. Different methods of statistical analysis can be used to process and analyze monitoring data. This paper presents the cluster analysis outcomes of the chemical composition of soil samples collected in the area of a copper mine during the annual monitoring. As a result of cluster analysis, all soil sampling points were divided into three clusters, which seem to be characterized by different mechanisms of pollution (background points without pollution, points in the sanitary protection zone of the quarry with aerogenic dust pollution, intermediate points with mixed pollution type). Based on the study outcomes, one can conclude that the application of cluster analysis in the soil monitoring data processing makes it possible to assess the boundaries of the influence zone of the dust pollution source with a little number of soil sampling points.

INTRODUCTION

Conducting of local environmental monitoring in the areas of development of mineral deposits is one of the important elements of ensuring the environmental safety of the territories. The need for such monitoring is stipulated in the environmental legislation of the Russian Federation and other countries [1].

Within the framework of the environmental monitoring program realization in the area of the Safianovskoye copper-pyrite ore deposit (Middle Ural, Russia), the Institute of Industrial Ecology (Ural Branch of the Russian Academy of Sciences, Yekaterinburg) conducts a long-term observation cycle on assessing soil contamination with metals that are specific for the ore and for the area where the deposit is located [2].

The impact on the soil is one of the important environmental aspects when carrying out mining operations. Monitoring of soil contamination provides information on the dynamics of accumulation of ore metals in the soil and makes it possible to assess the level of influence of the mining enterprise on the surrounding area [3].

Different methods of statistical analysis can be used to process and analyze monitoring data. This paper presents the results of a cluster analysis of the chemical composition of soil samples collected in the area of the copper mine during the annual monitoring.

METHODOLOGY AND RESULTS

Environmental monitoring works were started in 1992 as part of an environmental impact assessment procedure of the Safyanovskoye copper-pyrite deposit development project (Fig. 1). After the beginning of commercial mining of copper ore (1995), the studies of environmental components pollution, including soil, are conducted annually. A

Proceedings of the 44th International Conference on Applications of Mathematics in Engineering and Economics
AIP Conf. Proc. 2048, 060016-1–060016-5; https://doi.org/10.1063/1.5082131
Published by AIP Publishing. 978-0-7354-1774-8/$30.00

large amount of experimental data was accumulated, for the storing, processing and analyzing of which a special information system was created [4].

FIGURE 1. The area of Safyanovskoye copper-pyrite deposit.
(«SAS.Planet.Release.151111»)

In July 2017, soil samples were taken on the studied area. 12 samples of soil were selected in the zone of possible influence of the deposit. The layout of the sampling points is shown in Fig. 2. The points are located both in the sanitary protection zone of the enterprise with a size of 500 m, and at a distance of up to 5-6 km from the quarry and dumps of overburden.

FIGURE 2. The layout of sampling points.
(«SAS.Planet.Release.151111»)

The collected samples were analyzed to determine the following metals content: chromium, copper, nickel, zinc, arsenic, cadmium and lead. Copper and zinc are the main ore metals (the content is 0.5 - 2%); arsenic, lead and cadmium - ore micro components (the content is less than 0.01%); nickel can enter onto the territory from the enterprises located in the Rezh city (about 8 km from the deposit); chromium is an element that is characteristic of the geochemical background of the observations area.

The sampling and the chemical analysis of samples were carried out in accordance with the methods [5, 6, 7], as well as the requirements of state standards in Russia. The results of chemical analysis are presented in Table 1.

When assessing the impact of mining objects on the environment, an important task is to study the effect on the soil cover not only of the monitoring object itself, but also of other potential sources of pollution (such as industrial enterprises, roads, etc.) located in the study area. The information obtained makes it possible to determine the boundaries of the zone of influence of the studied object, as well as other possible sources. The mechanism of soil contamination at each specific point of selection may be different and depends on the point location relative to the objects of the mining enterprise. The main mechanism is an atmospheric transport of dust from the pit and dumps. In addition, soils can be contaminated with the sewage enriched by ore metals and micro components.

Table 1. The results of chemical analysis of soil samples in 2017, mg/kg.

No.	Cr	Cu	Zn	Ni	Pb	Cd	As
1	143,50	33,30	58,60	84,90	17,50	1,10	12,90
2	143,60	29,60	65,90	81,20	17,30	1,00	8,90
3	198,40	30,20	78,20	108,40	17,80	1,30	2,70
4	142,50	27,60	95,40	72,80	15,70	0,80	4,60
5	132,30	29,50	64,80	76,80	16,00	0,60	9,60
6	123,40	26,20	60,40	87,10	18,50	0,70	6,70
7	91,10	33,00	61,80	63,50	17,10	0,16	8,40
8	119,90	32,00	69,70	66,70	15,80	0,33	9,80
9	114,70	34,10	73,90	72,30	15,60	0,53	7,70
10	123,70	28,70	66,20	72,40	15,80	0,41	10,80
11	151,90	27,80	60,70	95,10	18,50	0,90	7,20
12	119,50	23,70	63,20	66,60	18,10	0,51	9,20
Mean	133,7	29,6	68,2	79,0	17,0	0,70	8,2

To interpret the obtained data and separate the observation points depending on their location, as well as possible soil contamination mechanisms, a cluster analysis of the results of the samples chemical analysis was carried out.

In general, cluster analysis is used to combine some objects into classes in such a way that the most similar objects fall into one class, and the objects of different classes differ as much as possible from each other. The quantitative measure of similarity is calculated by predetermined method taking into account the data characterizing the study objects.

All cluster algorithms need estimates of distances between clusters or objects, and it is clear that when calculating the distance, one needs to set the measurement scale. Since different elements are characterized by significantly different concentrations in soil (i.e., they have different scales), the analyzed data should be standardized before the analysis.

During the cluster analysis, the data were normalized to the maximum permissible concentrations of pollutants in soil [8, 9] and standardized. The standardized data for the analysis are shown in Table 2. The variables Var1-Var7 in Table 2 correspond to the following elements: Var1 - chromium, Var2 - copper, Var3 - zinc, Var4 - nickel, Var5 - lead, Var6 - cadmium and Var7 - arsenic.

Table 2. Standardized data for cluster analysis.

No.	Var1	Var2	Var3	Var4	Var5	Var6	Var7
1	0,37	1,18	-0,93	0,45	0,46	1,20	1,72
2	0,38	-0,01	-0,23	0,17	0,29	0,91	0,25
3	2,47	0,18	0,97	2,23	0,73	1,80	-2,02
4	0,34	-0,66	2,63	-0,47	-1,12	0,31	-1,32
5	-0,05	-0,05	-0,33	-0,17	-0,86	-0,28	0,51
6	-0,39	-1,11	-0,76	0,62	1,34	0,01	-0,55
7	-1,62	1,08	-0,62	-1,18	0,11	-1,59	0,07
8	-0,53	0,76	0,14	-0,93	-1,03	-1,08	0,58
9	-0,72	1,44	0,55	-0,51	-1,21	-0,49	-0,19
10	-0,38	-0,30	-0,20	-0,50	-1,03	-0,85	0,95
11	0,69	-0,59	-0,73	1,22	1,34	0,61	-0,37
12	-0,54	-1,91	-0,49	-0,94	0,99	-0,55	0,36

For cluster analysis, the standard package of Statistica (ver. 12) software was used. As the rule of association, Ward's method was used, and as a measure of proximity, the Euclidian distance. The proximity measure, determined by the Euclidean distance, is the geometric distance in n-dimensional space, which is calculated as follows (1):

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (1)$$

A dendrogram for 12 sampling points is shown in Fig. 3. Variables C_1 - C_12 on the dendrogram correspond to the sampling points 1 - 12.

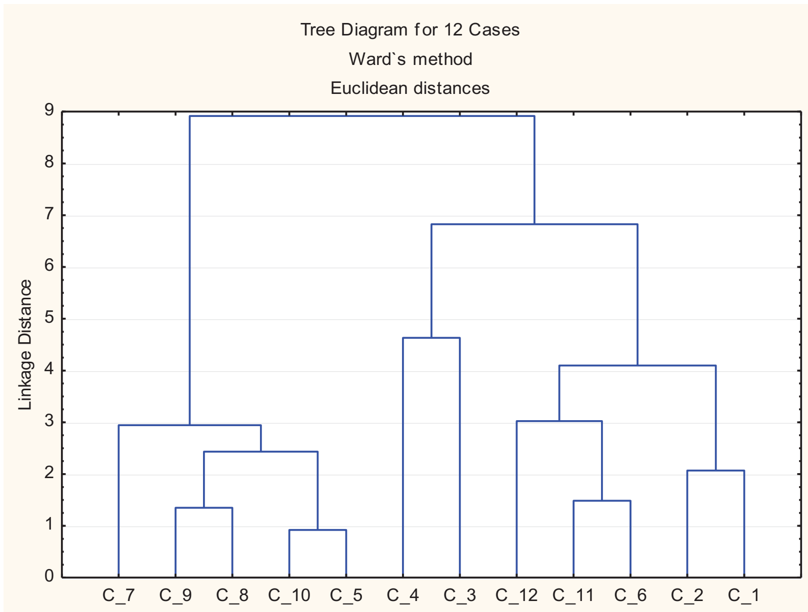

FIGURE 3. Dendrogram for 12 sampling points.

Based on the visual representation of the results, it can be assumed that the soil samples represent three natural clusters. To test this assumption, the initial data were divided by the method of K average values into 3 clusters, and the significance of the difference between the obtained groups was checked. The P-value for each variable was less than 0.05, which indicates the importance of this assumption.

The resulting clusters can be characterized as follows. The first class contains soil samples which were taken at points located far from the quarry's industrial site. The level of the studied elements content in soil at these points, apparently, is determined by natural geochemical conditions of the investigated region.

The second class contains soil samples taken at points located in the sanitary protection zone of the enterprise (up to 500 m from the boundary of the quarry and dumps). Pollution at these points occurs due to pollutants emission from the dumps and nearby technological roads as a result of pollutants sedimentation from the atmospheric air. These points allow estimating the spatial size and location of the quarry's influence zone on the territory.

The third class was also obtained which contains of two soil samples (3 and 4). One of the points (4) is located near the quarry and dumps, and the second (3) - at a distance of about 1 km. One can assume that the appearance of the third cluster is due to the fact that in these points a different pollution mechanism predominates. Near to the point 4 there is a temporary (seasonal) pond, which contains polluted waters from under the dumps. Point 3 is located on the bank of the stream, which begins nearby the industrial site of the quarry and is polluted by ore metals and micro components. Thus, one can suggest that the points 3 and 4 have a similar pollution mechanism, which is caused not only by atmospheric dust transport, but also by polluted waters from the quarry's area. That is that these points have a mixed pollution mechanism.

CONCLUSION

The paper presents some results of the interpretation of soil pollution monitoring data in the zone of possible influence of the operating copper mine (July 2017). Twelve soil samples were taken and their chemical analysis was carried out for the following contaminants: chromium, copper, nickel, zinc, arsenic, cadmium, mercury and lead. Then, a cluster analysis of the chemical analysis outcomes was carried out to separate the sampling points into groups, which are characterized by different distance from pollution source and possibly have other mechanism of pollution.

As a result of cluster analysis, all soil sampling points were divided into three clusters, which seem to be characterized by different mechanisms of pollution (background points without pollution, points in the sanitary protection zone of the quarry with aerogenic dust pollution, intermediate points with mixed pollution type - dust and polluted water). Based on the study outcomes, one can conclude that the application of cluster analysis in the soil monitoring data processing makes it possible to assess the boundaries of the influence zone of the dust pollution source with a little number of soil sampling points.

REFERENCES

[1] A. Medvedev, A. Sergeev and D. Melnichenka, "About Legislative Regulation of Local Environmental Monitoring in Russian Federation and Republic of Belarus", *15th International Multidisciplinary Scientific Geoconference SGEM 2015, Ecology, Economics, Education and Legislation, Conference Proceedings, Book 5, Vol. 2* (2015), p. 591-597.

[2] Derjagina S.E., Astafyeva O.V., Medvedev A.N., "Ecological monitoring as a tool for effective environmental management of the mining enterprise", Economics of Nature Use, No. 1, pp. 48-53 (2011), (in Russian).

[3] D.A. Tarasov, A.N. Medvedev, A.P. Sergeev, and A.G. Buevich, "Review and Possible Development Direction of the Methods for Modeling of Soil Pollutants Spatial Distribution", AIP Conference Proceedings 1863, 050014 (2017).

[4] M. Medvedev, "Information-Analytical System of Environmental Monitoring", *15th International Multidisciplinary Scientific Geoconference SGEM 2015, Conference Proceedings, Vol. I, Informatics, Geoinformatics, Photogrammetry and Remote Sensing* (2015), pp. 685-690.

[5] Vostroknutov G.A., *Temporary methodological guidance for conducting geochemical research in geoecological works* (PGO "Uralgeology", Ekaterinburg, 1991), (in Russian).

[6] GOST 17.4.3.01-83. Protection of Nature. Soil. General requirements for sampling (in Russian).

[7] GOST 17.4.4.02-84. Protection of Nature. Soil. Methods of selection and preparation of samples for chemical, bacteriological, helminthological analysis (in Russian).

[8] GN 2.1.7.2511-09. Approximate permissible concentration of chemical substances in soil (in Russian).

[9] GN 2.1.7.2041-06. The maximum permissible concentration of chemicals in soil (in Russian).